



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/728,721

12/05/2003

Todd D. Wakefield

5135

1732

64100

7590

04/03/2007

DANIEL P. MCCARTHY

P.O. BOX 71550

SALT LAKE CITY, UT 84171-0550

EXAMINER

SYED, FARHAN M

ART UNIT

PAPER NUMBER

2165

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
--	-----------	---------------

3 MONTHS

04/03/2007

PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No. 10/728,721	Applicant(s) WAKEFIELD ET AL.	
	Examiner Farhan M. Syed	Art Unit 2165	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 09 December 2006.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1,2,5-17,20-33 and 36-47 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1,2,5-17,20-33 and 36-47 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. Claims 1, 2, 5-17, 20-33, and 36-47 are pending.

Response to Remarks/Arguments

2. In the claims filed 09 December 2006, the Examiner acknowledges the cancellation of claims 3, 4, 18, 19, 34, and 35.
3. The terminal disclaimer filed on 09 December 2006 disclaiming the terminal portion of any patent granted on this application has been reviewed and is accepted. The terminal disclaimer has been recorded.
4. Applicant's arguments, see pages 12-13, filed 09 December 2006, with respect to the abstract have been fully considered and are persuasive. The objection of the non-Final office action mailed 08 June 2006 has been withdrawn.
5. Applicant's arguments filed 09 December 2006 have been fully considered but they are not persuasive for the reasons set forth below.

Applicant argues:

- (1) The prior art of record does not teach the limitations as described in claim 1.

The Examiner respectfully disagrees. Lemus teaches a computer program product located to one or more storage media devices usable to perform integration of

Art Unit: 2165

mixed format data (i.e. *"A system and method for generating structured data from unstructured or semi-structured data uses context-based natural language interpreters."* *"The systems and methods of the present invention can be used to extract information from text, and particularly from unstructured or short semi-structured messages, such as from email, pagers, or other communication devices."* The preceding text clearly indicates that a mixed format data are structured data from unstructured or semi-structured data. One or more storage media devices are pagers and other communication devices.)(Lemus, Abstract; see also paragraph [0007]), said computer program product comprising instructions executable by a computer to perform the functions of: accessing a database of data records, at least some of the data records containing both structured data and unstructured data (i.e. *"The system and method of the present invention have a number of aspects, including a system for receiving information in semi-structured or unstructured form from emails, pagers, and other communication methods, and converting that information into a structured form that can be usable in a database. The system and method also include methods for converting semi-structured data or unstructured data into a structured form suitable for use in a database. These methods can include the steps and processes described below or a subset of those steps and processes."* The preceding text clearly indicates that the prior art contains a database, where accessing a database of structured data is receiving information into a structured form that can be usable in a database. That is the process of receiving information encompasses accessing, because in order to receive information, one must access it. Furthermore, email messages are exemplified by the prior art to illustrate data records containing both structured data and unstructured data, where in an email message, there is a "To:" and "From:" field that clearly are structured fields and a "Body" that contains free text (i.e. unstructured data), all which are stored in a database. That is an email message is a semi-structured data that exemplifies the use of both structured data AND unstructured data.)(Page 1, paragraph [0009]; see also paragraphs [0014], [0016], [0029], and [0061]); the unstructured data of a particular data record including free text related to the structured data of that data record (i.e. *"A piece of*

Art Unit: 2165

text could contain one or more pieces of such semi-structured data. For instance, an email could detail, on separate lines, two rental availabilities. Each description of a rental availability would represent one piece of semi-structured data." The preceding text clearly illustrates that the email contains unstructured data, which are exemplified by separate lines that detail two rental availabilities.)(See also paragraphs [0016], [0029], and [0061]); using linguistic characteristics of the free text to extract relational facts from it, (i.e. "The systems and methods of the present invention can be used to extract information from text, and particularly from unstructured or short semi-structured messages, such as from email, pagers, or other communication devices." "For instance a field might consist of the data 123 456 7890, with that type of information being 'telephone number'" The preceding text clearly indicates that linguistic characteristics of free text is illustrated by '123 456 7890' and the relational facts extracted from it is that it exemplifies a 'telephone number.')(Page 1, paragraph [0007], [0013]), producing a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum being relatable to the structured data of the same data record (i.e. "Text pre-filtering attempts to perform data cleaning and massaging. The actual mechanisms used are dependent on the context. Atomization is the process of splitting a given piece of text with white spaces to get a list of the individual words." The process of atomization is producing a set of construed data reflecting at least one relational fact conveyed in the free text, where the atomization splits the given piece of text (i.e. free text) and extracting relational facts from it (i.e. telephone number, name, etc.) as addressed in the earlier limitation above.)(Page 3, paragraph [0031]); integrating the construed data with the particular structured data (i.e. "The resulting structured data can be used to create relational database records." The preceding text clearly indicates that the integrating the construed data is the creation of the relational database record that contains structured data.)(Lemus, Abstract); and rendering at least one visual representation of the integrated data (i.e. interface, 24, in figure 2.)(Figure 2).

Art Unit: 2165

Hence, the Applicant's arguments do not distinguish over the claimed invention over the prior art of record.

Any other arguments by the applicant are either more limiting than the claimed language or completely irrelevant.

Claim Objections

6. Claims 1, 2, 5-16, 20-32, and 36-47 are objected to because of the following informalities:

As per claim 1, lines 12-13 recite "producing a set of construed data reflecting at least one relational fact conveyed in the free text, each construed datum being relatable to the structured data of the same data record" is the same exact limitation, recited word for word as in claim 1, lines 14-15. The Examiner believes that the Applicant may have inadvertently duplicated the recited limitation and will treat the limitation as one limitation as oppose to two distinct limitations of claim 1.

As per claims 2, 5-16, 20-32, and 36-47, the word "The" should be used as oppose to "A," since the computer programming product has been defined in the respective independent claims. Appropriate correction is required.

Claim Rejections - 35 USC § 102

7. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

8. Claims 1, 2, 5-17, 20-33, and 36-47 are rejected under 35 U.S.C. 102(e) as being anticipated by Lemus (U.S. Patent Publication No. 2002/0156817 A1).

As per claims 1, 17, and 33, Lemus teaches a computer program product located to one or more storage media devices usable to perform integration of mixed format data (i.e. *"A system and method for generating structured data from unstructured or semi-structured data uses context-based natural language interpreters."* *"The systems and methods of the present invention can be used to extract information from text, and particularly from unstructured or short semi-structured messages, such as from email, pagers, or other communication devices."* The preceding text clearly indicates that a mixed format data are structured data from unstructured or semi-structured data. One or more storage media devices are pagers and other communication devices.)(Lemus, Abstract; see also paragraph [0007]), said computer program product comprising instructions executable by a computer to perform the functions of: accessing a database of data records, at least some of the data records containing both structured data and unstructured data (i.e. *"The system and method of the present invention have a number of aspects, including a system for receiving*

Art Unit: 2165

information in semi-structured or unstructured form from emails, pagers, and other communication methods, and converting that information into a structured form that can be usable in a database. The system and method also include methods for converting semi-structured data or unstructured data into a structured form suitable for use in a database. These methods can include the steps and processes described below or a subset of those steps and processes." The preceding text clearly indicates that the prior art contains a database, where accessing a database of structured data is receiving information into a structured form that can be usable in a database. That is the process of receiving information encompasses accessing, because in order to receive information, one must access it. Furthermore, email messages are exemplified by the prior art to illustrate data records containing both structured data and unstructured data, where in an email message, there is a "To:" and "From:" field that clearly are structured fields and a "Body" that contains free text (i.e. unstructured data), all which are stored in a database. That is an email message is a semi-structured data that exemplifies the use of both structured data AND unstructured data.)(Page 1, paragraph [0009]; see also paragraphs [0014], [0016], [0029], and [0061]); the unstructured data of a particular data record including free text related to the structured data of that data record (i.e. *"A piece of text could contain one or more pieces of such semi-structured data. For instance, an email could detail, on separate lines, two rental availabilities. Each description of a rental availability would represent one piece of semi-structured data."* The preceding text clearly illustrates that the email contains unstructured data, which are exemplified by separate lines that detail two rental availabilities.)(See also paragraphs [0016], [0029], and [0061]); using linguistic characteristics of the free text to extract relational facts from it, (i.e. *"The systems and methods of the present invention can be used to extract information from text, and particularly from unstructured or short semi-structured messages, such as from email, pagers, or other communication devices."* *"For instance a field might consist of the data 123 456 7890, with that type of information being 'telephone number'"* The preceding text clearly indicates that linguistic characteristics of free text is illustrated by '123 456 7890' and the relational facts extracted from it is that it exemplifies a 'telephone number.')(Page 1, paragraph [0007], [0013]), producing a set of construed data reflecting at

Art Unit: 2165

least one relational fact conveyed in the free text, each construed datum being relatable to the structured data of the same data record (i.e. *"Text pre-filtering attempts to perform data cleaning and massaging. The actual mechanisms used are dependent on the context. Atomization is the process of splitting a given piece of text with white spaces to get a list of the individual words."* The process of atomization is producing a set of construed data reflecting at least one relational fact conveyed in the free text, where the atomization splits the given piece of text (i.e. free text) and extracting relational facts from it (i.e. telephone number, name, etc.) as addressed in the earlier limitation above.)(Page 3, paragraph [0031]); integrating the construed data with the particular structured data (i.e. *"The resulting structured data can be used to create relational database records."* The preceding text clearly indicates that the integrating the construed data is the creation of the relational database record that contains structured data.)(Lemus, Abstract); and rendering at least one visual representation of the integrated data (i.e. interface, 24, in figure 2.)(Figure 2).

As per claim 2, Lemus teaches a computer program product wherein said extracting step uses caseframes (i.e. *"The system and method also include methods for converting semi-structured data or unstructured data into a structured form suitable for use in a database. These methods can include the steps and processes described below or a subset of those steps and processes."*)(Page 1, paragraph [0009]).

As per claims 5, 21, and 37, Lemus teaches a computer program product wherein said instructions are further executable to perform the function of producing a new database containing the integrated data produced by said integrating (i.e. *"Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database*

Art Unit: 2165

22. *The system includes an interface 24 to the means from which messages are received, such as over wireless application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses.*")(Page 3, paragraph [0061]).

As per claims 6, 22, and 38, Lemus teaches a computer program product wherein said instructions are further executable to perform the function of inserting the produced data into the database of structured data while performing said integrating the produced data (i.e. *"The system can thus receive data from one of a number of different sources and convert that data into structured data for use in a database, such as an Oracle or Sybase database. The resulting data can be used for data mining purposes. As a result, data entry can be fast and intuitive and can be flexible over one of a number of different devices. In addition, there is no need for the user to fill in structured fields and no need to learn complex input formats. As a result, there can be a reduction in data inconsistency and a significant elimination of re-keying, while allowing an entity that uses such a system to access and consolidate data that was previously scattered without impact on existing systems.*")(Page 3, paragraph [0063]).

As per claims 7, 23, and 39, Lemus teaches a computer program product wherein said instructions are further executable to perform the function of creating a new database while performing said integrating the produced data (i.e. *"Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database 22. The system includes an interface 24 to the means from which messages are received, such as over wireless*

Art Unit: 2165

application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses.”(Page 3, paragraph [0061]).

As per claims 8, 24, and 40, Lemus teaches a computer program product wherein the instructions are further executable to produce a new relational database containing the integrated data produced by said integrating (i.e. *“Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database 22. The system includes an interface 24 to the means from which messages are received, such as over wireless application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses.” “The system can thus receive data from one of a number of different sources and convert that data into structured data for use in a database, such as an Oracle or Sybase database. The resulting data can be used for data mining purposes. As a result, data entry can be fast and intuitive and can be flexible over one of a number of different devices. In addition, there is no need for the user to fill in structured fields and no need to learn complex input formats. As a result, there can be a reduction in data inconsistency and a significant elimination of re-keying, while allowing an entity that uses such a system to access and consolidate data that was previously scattered without impact on existing systems.”(Page 3, paragraphs [0061] and [0063]).*

As per claims 9, 25, and 41, Lemus teaches a computer program product wherein the instructions are further executable to produce a file containing the integrated data produced by said integrating (i.e. "Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database 22. The system includes an interface 24 to the means from which messages are received, such as over wireless application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses.")(Page 3, paragraph [0061]).

As per claims 10, 26, and 42, Lemus teaches a computer program product wherein the instructions are further executable to produce a file having a format selected from the group of XML, character separated values, spreadsheet formats and file-based database structures (i.e. "Data is defined to be a string of symbols, which may be chosen, for example, from the UNICODE character set. In the preferred embodiment the symbols are strings in the language Perl and are semi-structured, although the present invention could work with unstructured data. The term semi-structured data (SSD) is used in the manner described in the article entitled "Learning Information Extraction Rules for Semi-structured and Free Text," by Stephen Soderland, Machine Learning, 1-44 (this definition is followed rather than the definition used by the database community, which refers to this as "structured text"). SSD is generally somewhere between data in a rigidly specified grammar (such as XML or HTML) and free text in languages such as English. Typically SSD possesses almost no grammar, and is very telegraphic in style. Examples of SSD may be drawn from classified advertisements in newspapers, such as: 1 Earl's Court, SW5, the rent is \$40 per

Art Unit: 2165

week. "Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database 22. The system includes an interface 24 to the means from which messages are received, such as over wireless application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses."(Page 1, paragraph [0014], [0061]).

As per claims 11 and 27, Lemus teaches a computer program product further comprising: a processing unit coupled to said one or more storage media devices, said processing unit being capable of executing said instructions (i.e. "The computer system that implements the steps and processes described above can be or include application specific integrated circuits (ASICs) or can include one or more personal computers, servers, or other such computational devices or group of devices.")(page 3, paragraph [0062]); and an execution command unit, whereby operation of said instructions and said processing unit may be commanded or controlled (i.e. "The system can thus receive data from one of a number of different sources and convert that data into structured data for use in a database, such as an Oracle or Sybase database. The resulting data can be used for data mining purposes. As a result, data entry can be fast and intuitive and can be flexible over one of a number of different devices. In addition, there is no need for the user to fill in structured fields and no need to learn complex input formats. As a result, there can be a reduction in data inconsistency and a significant elimination of re-keying, while allowing an entity that uses such a system to access and consolidate data that was previously scattered without impact on existing systems.")(page 3, paragraph [0063]).

As per claims 12, 28, and 43, Lemus teaches a computer program product wherein said instructions are further executable to combine similar attributes for the extracted relational facts produced in performing said extracting relational facts from the free text (i.e. "Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database 22. The system includes an interface 24 to the means from which messages are received, such as over wireless application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses." "The system can thus receive data from one of a number of different sources and convert that data into structured data for use in a database, such as an Oracle or Sybase database. The resulting data can be used for data mining purposes. As a result, data entry can be fast and intuitive and can be flexible over one of a number of different devices. In addition, there is no need for the user to fill in structured fields and no need to learn complex input formats. As a result, there can be a reduction in data inconsistency and a significant elimination of re-keying, while allowing an entity that uses such a system to access and consolidate data that was previously scattered without impact on existing systems.") (Page 3, paragraphs [0061] and [0063]).

As per claims 13, 29, and 44, Lemus teaches a computer program product wherein said instructions are further executable to combine similar relation types for the extracted relational facts produced in performing said extracting relational facts from the free text (i.e. "Referring to FIG. 2, the system of the present invention can be implemented on one or more special purpose or general purpose computers 20, appropriately configured and/or programmed, and coupled to a database 22. The system includes an interface 24 to the means from which messages

Art Unit: 2165

are received, such as over wireless application protocol (WAP), short message service (SMS), email 26, pager 28, document 30, or voice recognition system 32; and an interface 34 to database 22 into which the data is stored in fields. The input can be in text or in a publicly available proprietary form, such as a word processor or PDF document. The data in the database can then be used for searching, report generation, business process management, or other uses." "The system can thus receive data from one of a number of different sources and convert that data into structured data for use in a database, such as an Oracle or Sybase database. The resulting data can be used for data mining purposes. As a result, data entry can be fast and intuitive and can be flexible over one of a number of different devices. In addition, there is no need for the user to fill in structured fields and no need to learn complex input formats. As a result, there can be a reduction in data inconsistency and a significant elimination of re-keying, while allowing an entity that uses such a system to access and consolidate data that was previously scattered without impact on existing systems.")(Page 3, paragraphs [0061] and [0063]).

As per claims 14, 30, and 45, Lemus teaches a computer program product wherein said instructions provide relational facts with domain roles applied in performing said extracting relational facts from the free text (i.e. *"Domain-dependent processing is the manipulation of parts of the semi-structured document that is dependent on the domain of discourse that the information resides in. For example, semantic information peculiar to the domain of discourse may be used to identify terms and present them in a normalized form. If the domain relates to motorcars, this semantic context may identify terms such as "VW" and "Volksy" and represent them both of them as the normal term "Volkswagen." The extraction engine provides facilities to accomplish such manipulations. These manipulations consist of term rewrites that utilize lexicons. The triggering of manipulations often relies on the use of the intelligence services described below."*)(Page 4, paragraph 0073).

As per claims 15, 31, and 46, Lemus teaches a computer program product wherein said instructions store the relational facts produced in performing said extracting relational facts from the free text (i.e. "Text pre-filtering attempts to perform data cleaning and massaging. The actual mechanisms used are dependent on the context. Atomization is the process of splitting a given piece of text with white spaces to get a list of the individual words." "This step uses a set of pattern matching and replace expressions, which insert white space in correct places, using basic syntactic typing rules. The rules are context dependent, and in the preferred embodiment are stored in a separate database called, for example, "AtomizeRules." The rules can be programmed in Perl or other language that supports regular expressions and string manipulation. A rule is a regular expression that specifies how white space is to be inserted. For instance, the example might contain a regular expression whose purpose is to insert white space before commas and full stops." "In some instances, it is desirable for several words to be treated as a single atom because they represent a single semantic entity. This step handles these cases. A separate database table contains patterns, including white spaces, which are to be replaced with the same words but with the white space replaced with an underscore. In the preferred embodiment, only atoms from the previous stage are used, and are combined into a piece of text again (with only one space between each atom), and apply the expressions found in the database." "The fields of the record corresponding to the context are populated with the classified atoms and/or atom sequences; i.e., a context may include several types of information, such as name, city, and state, and the atoms are classified into those types.") (page 2, paragraphs [0031] and [0039]; page 3, paragraph [0049], [0057]).

As per claims 16, 32, and 47, Lemus teaches a computer program product wherein the extracted relational facts produced in performing said extracting relational facts and the integrated data produced by the performance of said integrating the produced data includes reference information to the original free text (i.e. "The systems and

methods of the present invention can be used to extract information from text, and particularly from unstructured or short semi-structured messages, such as from email, pagers, or other communication devices." "Text pre-filtering attempts to perform data cleaning and massaging. The actual mechanisms used are dependent on the context. Atomization is the process of splitting a given piece of text with white spaces to get a list of the individual words." "This step uses a set of pattern matching and replace expressions, which insert white space in correct places, using basic syntactic typing rules. The rules are context dependent, and in the preferred embodiment are stored in a separate database called, for example, "AtomizeRules." The rules can be programmed in Perl or other language that supports regular expressions and string manipulation. A rule is a regular expression that specifies how white space is to be inserted. For instance, the example might contain a regular expression whose purpose is to insert white space before commas and full stops." "In some instances, it is desirable for several words to be treated as a single atom because they represent a single semantic entity. This step handles these cases. A separate database table contains patterns, including white spaces, which are to be replaced with the same words but with the white space replaced with an underscore. In the preferred embodiment, only atoms from the previous stage are used, and are combined into a piece of text again (with only one space between each atom), and apply the expressions found in the database." "The fields of the record corresponding to the context are populated with the classified atoms and/or atom sequences; i.e., a context may include several types of information, such as name, city, and state, and the atoms are classified into those types."(page 1, paragraph [0007]; page 2, paragraphs [0031] and [0039]; page 3, paragraph [0049], [0057]).

As per claims 20 and 36, Lemus teaches a computer program product wherein said instructions are further executable to perform the function of applying caseframes while performing said interpreting the free text (i.e. *"Initially, the text file is context-classified as an information source for one or several data structures. The context is the surrounding information that identifies the characteristics of the information available in the text file."* "Context identification classifies

Art Unit: 2165

the textual data according to a predefined or user-defined context. Context identification might be made using one or more of the following methods: (1) User classification, (2) Automatic classification via keyword identification, (3) Automatic classification via data-origin or data-destination, and (4) Automatic classification via pattern identification, such as with machine learning techniques.") (page 2, paragraph [0023] and [0024]).

Conclusion

9. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the mailing date of this final action.

Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Farhan M. Syed whose telephone number is 571-272-7191. The examiner can normally be reached on 8:30AM-5:00 PM.

Art Unit: 2165

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jeffrey Gaffin can be reached on 571-272-4146. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

FMS


JEFFREY GAFFIN
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100